

AD-A127 976

THE AFEEES-NLS (ARMED FORCES ENTRANCE EXAMINATION
STATION-NATIONAL LONGITU. (U) RAND CORP SANTA MONICA CA
J R HOSEK ET AL. JAN 83 RAND/N-1930-MRAL

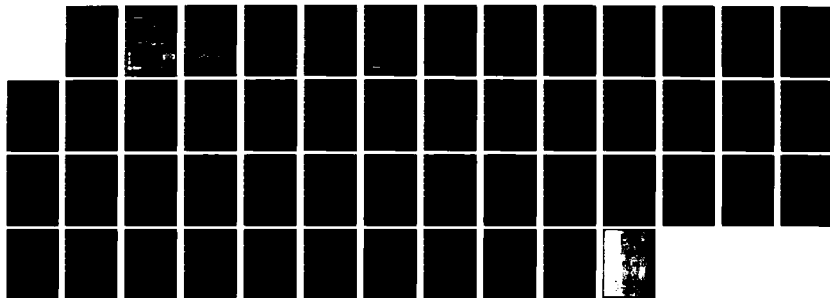
1/1

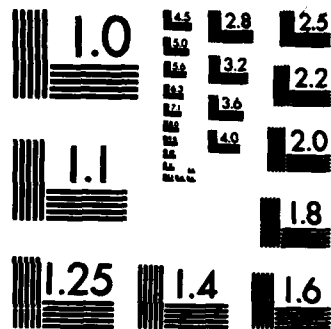
UNCLASSIFIED

MDA903-80-C-0652

F/G 5/9

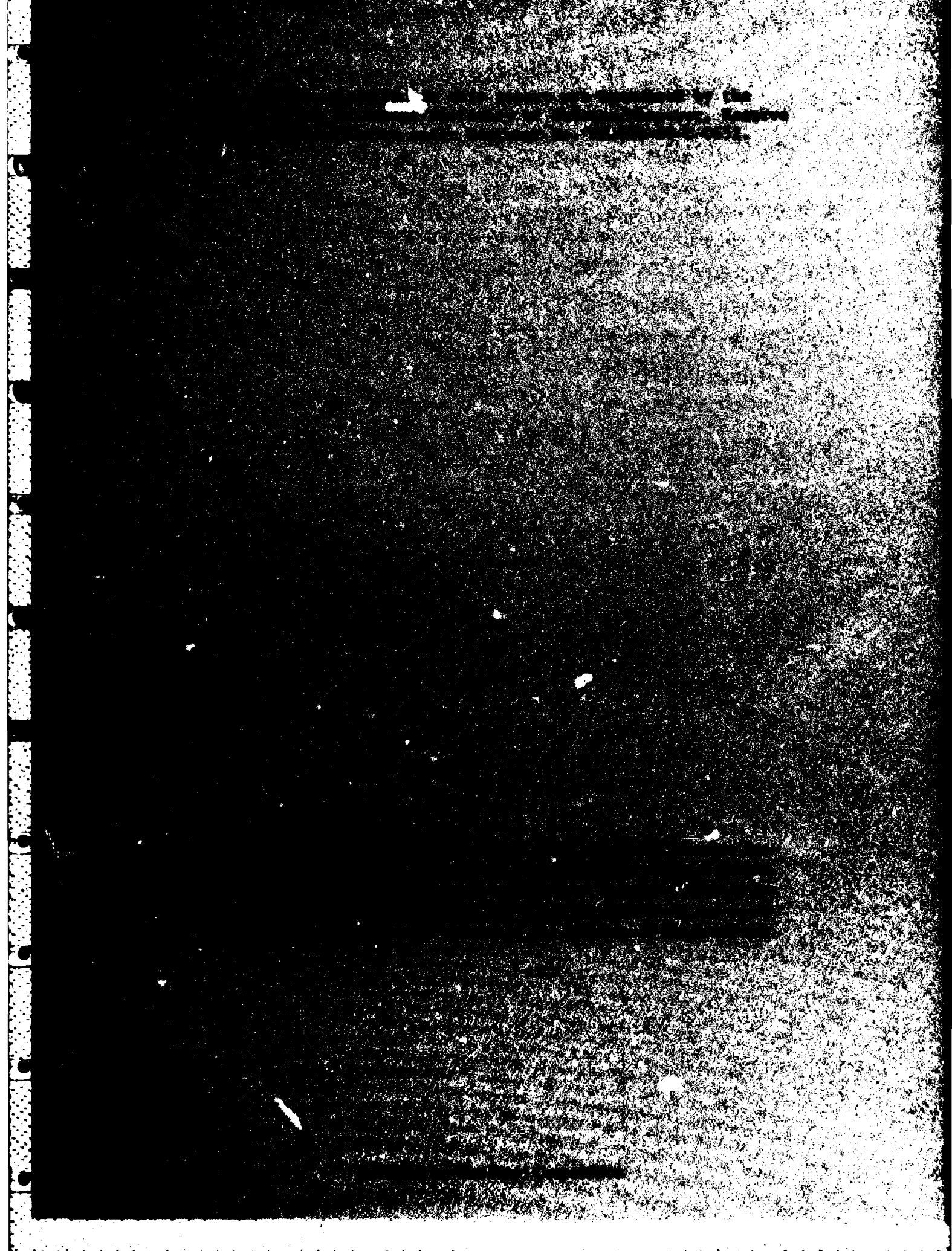
NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

ADA 127976



UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|--|-------------------------------------|--|
| 1. REPORT NUMBER N-1930-MRAL | 2. GOVT ACCESSION NO. 40 4129 72 | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle) The AFES-NLS Database: A Choice-Based Sample for Studying Enlistment and Post-Enlistment Outcomes | | 5. TYPE OF REPORT & PERIOD COVERED Interim |
| | | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s) James R. Hosek, Christine E. Peterson | | 8. CONTRACT OR GRANT NUMBER(s) MDA903-80-C-0652 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS The Rand Corporation 1700 Main Street Santa Monica, CA. 90406 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
| 11. CONTROLLING OFFICE NAME AND ADDRESS Assistant Secretary of Defense Manpower, Reserve Affairs & Logistics Washington, D.C. 20301 | | 12. REPORT DATE January 1983 |
| | | 13. NUMBER OF PAGES 38 |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) | | 15. SECURITY CLASS. (of this report) UNCLASSIFIED |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |
| 16. DISTRIBUTION STATEMENT (of this Report) Approved for Public Release; Distribution Unlimited | | |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) No Restrictions | | |
| 18. SUPPLEMENTARY NOTES | | |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Enlisted Personnel Recruiting Surveys | | |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number) See Reverse Side | | |

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Describes prospective research and policy applications of a newly constructed database intended for the analysis of enlistment decisionmaking and post-enlistment outcomes. The database expands the potential for analysis beyond that inherent in aggregate data, the form of data used in nearly all existing studies of enlistments. The constructed database should prove useful in analyzing the determinants of enlistment for different groups of young people, such as students and nonstudents, employed and not employed, and those in their late teens and early twenties. Knowledge of these determinants should be applicable to the formulation of recruiting strategies, to the design of enlistment incentives, and to the prediction of recruit quantity and quality under different economic conditions and accession policies.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

A RAND NOTE

THE AFEES-NLS DATABASE: A CHOICE-BASED SAMPLE
FOR STUDYING ENLISTMENT AND POST-ENLISTMENT
OUTCOMES

James R. Hosek, Christine E. Peterson

January 1983

N-1930-MRAL

Prepared for

The Office of the Assistant Secretary of
Defense/Manpower, Reserve Affairs and
Logistics

35th
Year



PREFACE

This Note describes prospective research and policy applications of a newly constructed database intended for the analysis of enlistment decisionmaking and post-enlistment outcomes. The database expands the potential for analysis beyond that inherent in aggregate data, the form of data used in nearly all existing studies of enlistments. The constructed database should prove useful in analyzing the determinants of enlistment for different groups of young people, such as students and nonstudents, employed and not employed, and those in their late teens and early twenties. Knowledge of these determinants should be applicable to the formulation of recruiting strategies, to the design of enlistment incentives, and to the prediction of recruit quantity and quality under different economic conditions and accession policies.

The Note was prepared for the Office of the Assistant Secretary of Defense (Manpower, Reserve Affairs, and Logistics) under Contract No. MDA903-80-C-0652. The research is part of Rand's Manpower, Mobilization, and Readiness Program.



| | |
|-------------------|--|
| Accession For | |
| NTIS GRA&I | <input checked="checked" type="checkbox"/> |
| DTIC TAB | <input type="checkbox"/> |
| Unannounced | <input type="checkbox"/> |
| Justification | |
| By | |
| Distribution/ | |
| Availability Code | |
| ist | Avail and/or Special |
| A | |

SUMMARY

Different kinds of data and models are useful for different purposes. Most research on enlistment behavior uses aggregate data, the primary purpose being to forecast future enlistments. The relatively few aggregate-level explanatory variables in these models make them expedient for forecasting group enlistments. But aggregate variables, especially if few in number, are blunt instruments for probing the determinants of individual enlistment behavior and subsequent outcomes. To study the latter, data on the individuals themselves are most appropriate.

The microanalysis of enlistment has been stymied by a lack of data. To overcome this problem, data from two different samples have been pooled to form a "choice-based" sample of enlistees and nonenlistees. Data on enlistees come from the spring wave of the 1979 DoD Survey of Personnel Entering Military Service [AFEES (Armed Forces Entrance Examination Station) Survey], and data on nonenlistees come from the 1979 wave of the National Longitudinal Survey of Youth (NLS survey). The surveys were given at approximately the same time and contain many comparable variables. The pooled observations from the AFEES and NLS surveys have been augmented by data drawn from additional sources. All together, the data comprise the AFEES-NLS database. Appropriately weighted, data in such a choice-based sample have the desirable property of being able to yield statistically consistent, efficient estimates of multivariate models. In other words, the AFEES-NLS choice-based sample data can be used to fit regression models of enlistment choice.

Several reasons justify the use of microdata, specifically the AFEES-NLS database, to study enlistment behavior and post-enlistment outcomes. First, because microdata frequently contain a wide variety of information about the individual, his socioeconomic background, and his local economic circumstances, a rich array of hypotheses can be tested. Second, because the data pertain to the individual himself, microdata allow the analyst to study how changes in these data, rather than in some aggregate surrogate for these data, relate to the individual's

enlistment choice. For example, microdata specify the individual's wage rate, not the average wage of youth or the average wage in manufacturing. Third, aggregate data on enlistments include only those persons who have chosen to enlist; they omit nonenlistees. In contrast, the microdata in the AFEES-NLS database include both enlistees and nonenlistees. The microdata can help predict who among the nonenlistees is more or less likely to enlist as the conditions affecting the enlistment decision change. Such conditions include not only education, wage, and employment status, but also military pay, enlistment incentives, and recruiting efforts. Fourth, the broad range of information available in the AFEES-NLS database relative to aggregate data creates a potential for analyzing which individual characteristics, in addition to the usual ones--education and AFQT, are useful predictors of first-term performance (e.g., attrition, promotion, reenlistment) among enlistees. Fifth, microdata permit great flexibility in studying different segments of the population eligible to enlist. For instance, separate models can be estimated for high school seniors, high school dropouts, college attendees, and persons not in school. Statistical estimates can therefore confirm or reject notions about whether different groups have different propensities to enlist or whether they respond differently to changes in variables like wage rate, employment status, or military/civilian pay. By implication, the data can paint fairly precise portraits of the enlistment behavior of various groups under various conditions. Analyses based on aggregate data have not, and cannot, exhibit this flexibility.

The AFEES-NLS database can be employed to fit alternative models of enlistment behavior. One kind of model describes whether or not to enlist, and, more generally, which service to enter. A variant would describe which occupational area to enlist in. A second kind of model relates enlistment intentions to actual enlistment behavior. This kind of model is useful for analyzing and improving the capability of market surveys of enlistment intentions to predict who will eventually enter the services. A third kind of model treats enlistment behavior along with post-enlistment outcomes that include training attrition, post-training attrition, promotion, and reenlistment. These models can be estimated because the AFEES-NLS database is longitudinal in nature:

data on individuals will accumulate as they are followed over time. Models that consider enlistment and post-enlistment behavior as a system of equations may prove to be especially useful in predicting first-term performance and in validating or improving enlistment standards. An analogue to the third kind of model concerns the behavior of persons who choose not to enlist. In particular, the employment and earnings experience of nonenlistees can be studied and compared with that of enlistees.

ACKNOWLEDGMENTS

Dr. G. Thomas Sicilia, Director, Accession Policy (MRA&L), offered support and encouragement throughout this research. Important portions of file creation were undertaken by the Defense Manpower Data Center (Monterey), and the authors would especially like to thank Jane Crotser and Robbie Brandewie in this regard. Paul Nickens, Recruit Market Network Branch, Defense Manpower Data Center (Alexandria) kindly supplied the information on recruiters. Zahava D. Doering, Chief, Survey and Market Analysis Division (DMDC), and Lt. Col. P. A. Lerro (AP) offered perceptive suggestions in their roles as project monitors. At Rand, Jacqueline McGee gave valuable help in locating supplementary data sources for the AFEES-NLS database, and Larry Painter programmed the comparability code for combining the AFEES and NLS data sets. Richard Buddin and Nat Schenker helped in the painstaking process of defining variables, and Judy Fernandez provided a useful review of the Note. Jeanne Heller and Joyce Peterson considerably improved the Note's structure and clarity, and Lorraine Scruggs supplied excellent secretarial support.

CONTENTS

| | |
|-----------------------|------|
| PREFACE | iii |
| SUMMARY | v |
| ACKNOWLEDGMENTS | ix |
| TABLES | xiii |

Section

| | |
|---|----|
| I. INTRODUCTION | 1 |
| II. THE MICRODATA FILE | 4 |
| A Choice-Based Sample | 4 |
| Creating the AFEES-NLS Database | 5 |
| Flexibility of the File | 6 |
| III. USING THE DATA | 10 |
| Advantages of Microdata | 10 |
| Models of Enlistment Choice | 12 |
| IV. POLICY APPLICATIONS AND RESEARCH TOPICS | 15 |
| Policy Applications | 15 |
| Research Topics the Data Support | 16 |
| Some Enlistment Hypotheses | 17 |
| V. CONCLUSIONS | 20 |

Appendix

| | |
|--|----|
| CREATION OF THE AFEES-NLS DATABASE | 21 |
|--|----|

TABLES

| | |
|---|----|
| 1. Analysis Variables on the 1979 AFEES-NLS Database | 7 |
| A.1. Correspondence of Interview Dates between 1979 AFEES and 1979 NLS | 22 |
| A.2. Comparable Variables between 1979 AFEES and 1979 NLS Surveys | 23 |

I. INTRODUCTION

Defense manpower analysts need to know not only how various economic and sociodemographic trends affect enlistment, but also how individuals reach the decision to enlist. Without this information, the Office of the Secretary of Defense and the services can predict and respond to fluctuations in the quality and quantity of enlistments, but they will lack precise and detailed information about what factors, including those susceptible to policy tools, systematically influence them. Microdata on enlistment behavior can complement aggregate data in providing this information.

Especially since the advent of the All-Volunteer Force (AVF) in 1973, it has been vital for the services to forecast accurately the number of new enlistees they can expect over the next five to ten years. Thus most analyses of enlistments have been directed toward the goal of forecasting. In this work, aggregate data on enlistments (or on accessions) have been statistically related to routinely collected, aggregate variables such as the size of the young male population, the unemployment rate, the number of recruiters, and the ratio of military wage to average wage in manufacturing. These models have estimated enlistments for all the services, individually and collectively; they have particularly focused on predicting male "high-quality" enlistments--high school graduates who score in the upper half of the Armed Forces Qualification Test (AFQT). The models have been valuable in predicting whether, under foreseeable military and civilian economic conditions, the services could meet their recruiting objectives.

Despite their effectiveness in predicting recruitment trends, the aggregate (or macro) models are limited in their insights into how the services might influence enlistment decisions. The promise of developing a microdata base, then, lies in expanding the capability to analyze enlistment behavior, both its determinants and consequences. Unlike aggregate data, microdata can describe an individual's specific characteristics and circumstances, for example, his sociodemographic background, labor force experience, wage, employment status, student

status, and education aspirations. These data can help show which factors influence the decision to enlist, whether the influence varies over different groups of young people, and whether the factors are related to post-enlistment outcomes such as attrition and reenlistment. Such information may be a useful input into the continuing task of improving enlistment incentives, recruitment strategies, and accession goals. To realize this potential, the first step is to create a microdata file for the study of enlistment.

The absence of enlistment analyses based on microdata can almost surely be attributed to a virtual absence of data rather than a lack of interest in analyzing it. Chief among the reasons for the lack of microdata are two facts: samples of individuals are typically random (or stratified random), and only a small fraction of the relevant population enlists at a given time. As a result of the latter, an analyst would need an extremely large random sample to get reliable information about enlistment decisions. Existing samples--the Current Population Survey, the 1967 National Longitudinal Survey (NLS) of Young Men and Young Women, and others--have not been large enough to offer more than a few hundred observations on enlistments in a given year. While that helps, such small samples hinder thorough statistical analysis. In addition, to study decisionmaking, data must consist of people who face the same set of alternatives under the same set of conditions. This temporal aspect precluded our using the enlisted population in the 1979 NLS because they joined prior to and not during the survey period. Creating a "choice-based" sample can circumvent these problems. "Choice-based" means a sample containing data on individuals grouped according to the alternative they have chosen in a decision, under similar conditions, in the same period of time. We have been able to create a choice-based sample through merging two existing surveys on enlistees and nonenlistees. Thereby, the delay and expense of undertaking a sufficiently large random sample have been avoided. We call our file the AFEES-NLS database. As we explain in Section II, by weighting these data appropriately, we can produce statistically consistent and efficient estimates of population distributions and

responses, despite the fact that the sample groups are not represented in the same proportions as in the nation's youth population.[1]

The AFEES-NLS file makes available, for the first time in a single database, enough observations on individual enlistees and nonenlistees to support statistical analyses of enlistment behavior among different groups of youth. Analysis based on this material should overcome many of the shortcomings of enlistment analysis based on macrodata. Further, the methodology used to create the database provides a replicable prototype. Thus, it could be used to construct series of microdata samples that can be designed for special-purpose analyses, compared with one another or used to complement the enlistment data that the services routinely collect. An especially attractive feature of the methodology is its economy.

In the next section, we characterize the data sources, define "choice-based" samples more fully, and briefly describe how the microdata file was created. The Appendix describes in greater detail how the database was created. Section III explains the technical advantages of microdata for manpower policy analysis and describes some analytical models the microdata file supports. In Section IV, we suggest some policy areas, research topics, and enlistment hypotheses that analysis of the microdata file can illuminate. Finally, Section V contains the conclusions.

[1] See Charles F. Manski and Steven R. Lerman, "The Estimation of Choice Probabilities from Choice-Based Samples," Econometrica, Vol. 45, No. 8, November 1977, pp. 1977-1988.

II. THE MICRODATA FILE

A CHOICE-BASED SAMPLE

Fundamentally, the AFEES-NLS database comprises observations on individuals drawn from two separate surveys. Observations on enlistees come from the spring wave of the 1979 DoD Survey of Personnel Entering Military Service (1979 AFEES Survey) that surveyed individuals at Armed Forces Entrance Examination Stations just after they signed an enlistment contract to serve on active duty. The second, or fall, wave was not used because interview dates did not overlap with the 1979 NLS. The AFEES Survey was the product of a Rand-DoD project.[1] Observations on nonenlistees come from the first wave of the new National Longitudinal Survey of Labor Force Behavior, Youth Survey (1979 NLS), which is a stratified random sample of young men and women aged 14 to 22 in 1979. The Department of Labor contracted with the Center for Human Resource Research (Ohio State University) for a new longitudinal study on the labor market experiences of young men and women that began in 1979. In addition, the Department of Defense and the services provided funds to include a sample of active duty personnel among the interviewees. The NLS Youth Survey is administered by the National Opinion Research Center (University of Chicago) under a subcontract from the Center for Human Resource Research.[2]

Put together, the common data elements of the two surveys form a "choice-based" sample, which we call the AFEES-NLS. This sample contains data on individuals identified as having made or not made a

[1] For more detail, see Z. D. Doering, D. W. Grissmer, and J. Morse, 1979 DoD Survey of Personnel Entering Military Service: Wave 1 User's Manual and Codebook, The Rand Corporation, N-1605-MRAL, November 1980.

[2] For more detail, see National Longitudinal Survey of Labor Force Behavior, Youth Survey 1979, published by the National Opinion Research Center, University of Chicago, and the Center for Human Resource Research, Ohio State University, NORC-4270 1/79 OMB 44 R-1671. See also The National Longitudinal Surveys Handbook 1979, Center for Human Resource Research, College of Administrative Science, The Ohio State University.

choice to enlist, under similar circumstances, in the same period of time. In general, choice-based sampling differs from random sampling in that the rate at which the population is sampled is allowed to depend on the choice that has been made, whereas random sampling does not vary the sampling rate according to choice. Usually, infrequently occurring outcomes are sampled at higher rates in choice-based sampling, and in our case the infrequent outcome is enlistment. In the population, less than 10 percent of young men enlist in a given year, but the AFEES-NLS choice-based sample contains over 50 percent enlistees. Nevertheless, when appropriately weighted, the observations in the sample are representative of the populations of enlistees and nonenlistees. The weighting reflects the proportion of individuals in the sample who enlisted relative to the proportion in the population who enlisted. Moreover, when appropriately weighted, choice-based sample data will produce statistically consistent estimates of the coefficients in multivariate models of choice. Thus choice-based samples possess much the same capability for statistical analyses of choice as do random samples, but choice-based samples have the advantage of being potentially cheaper. It is not surprising that choice-based sampling is now coming into wide use.

CREATING THE AFEES-NLS DATABASE

There are two criteria for creating a choice-based sample by pooling separate data files: the information must be comparable temporally and substantively. Our microdata satisfy these criteria. The AFEES and the NLS surveys were administered in spring 1979 and both focused on demographics, education, and employment histories. Both also included geographic indicators that allowed us to add data on local employment conditions, youth population, recruiter population, and availability of enlistment options--variables that can influence enlistment decisions and post-enlistment outcomes. The congruence in timing and coverage of the AFEES and NLS surveys stems from DoD's involvement in both surveys. The sources of these additional data are described in the Appendix.

To make sure that the AFEES and NLS files were substantively comparable, we recognized and adjusted for certain differences. First, we deleted individuals in the NLS survey who were deemed not eligible for enlistment in 1979. Deletions include persons under age 17 and persons who reported that a health condition would impede their ability to work. Second, the NLS file contained many more variables than the AFEES. Hence, in pooling the files, we kept only the subset of NLS variables that were comparable to those in the AFEES or could be used to create comparable variables. However, the deleted observations and variables can be retrieved should the need arise. Third, because the AFEES and NLS employed different questionnaires, it required considerable effort to ensure that the retained variables were strictly comparable. The surveys asked similar types of questions, but they were by no means identical:

- o The AFEES asked simpler, more direct questions than the NLS.
- o The surveys used different skip patterns.
- o Differences in emphasis made some apparently similar variables in the two surveys quite different.
- o The surveys had different objectives in asking particular questions.
- o The NLS was by personal interview, the AFEES survey was self-administered.

We could identify problems and correct errors because we could check with the basic file; the Appendix describes this process in detail. Table 1 displays variables on the AFEES-NLS database which can be used for analyzing enlistment decisionmaking.

FLEXIBILITY OF THE FILE

The AFEES-NLS database is not a static file. We can merge in additional geographic-based data as we discover new sources of information pertinent to our analysis. For example, we have added monthly deviations from trends in state employment rates that became available after the file was formed,[3] and 1980 Census data can be

[3] The data were developed in a Rand project on forecasting accessions and are described in the Appendix.

Table 1

ANALYSIS VARIABLES ON THE 1979 AFEES-NLS DATABASE

- I. Analysis Variables on the 1979 AFEES and 1979 NLS surveys
 1. Enlistment Status
 - Enlist/not enlist
 - Service chosen
 - Occupational area entered in service
 - Length of term
 2. Age
 3. Race
 - White
 - Black
 - Hispanic
 4. Education
 - Highest grade completed
 - Full-time student
 - Part-time student
 - More education desired
 - Eldest sibling more educated
 5. Armed Services Vocational Aptitude Battery scores
 6. AFQT percentile
 7. Marital status
 - Single/not single
 8. Socioeconomic background
 - Father's education
 - Mother's education
 - Family income (1978)
 - No siblings
 - Number of siblings
 - Lived with both parents at age 14
 9. Labor force experience last year (1978)
 - Never worked
 - Unemployed all year
 - Full year - full-time
 - Full year - part-time
 - Part year - full-time
 - Part year - part-time
 10. Earnings experience
 - Personal income (1978)
 - Wage at current or most recent job
 11. Current employment experience
 - Not employed/employed
 - Tenure, if employed
 - Duration not employed, if not employed
 12. Location of residence
 - County for NLS observations
 - ZIP for AFEES observations; mapped to county
 - Region (4 or 9 census divisions)

Table 1 - continued

ANALYSIS VARIABLES ON THE 1979 AFEES-NLS DATABASE

II. Variables Merged into AFEES and NLS Records at the County Level

1. Number of recruiters by service (spring 1979)
2. Number of high school graduates by sex (1977 data)
3. Number of high school seniors by sex (1978 data)
4. Population counts by sex and race (white/black-minority)
for age groups 10-14, 15-19, 20-24, 25-29 (1978 estimates)
5. Employment levels by one-digit SIC (Standard Industrial
Classification) industry, annually, for the years 1969 to 1977
6. (1980 census data can be merged in when available)

III. Variables Merged into AFEES and NLS Records at the Armed Forces
Entrance Examination Station Level

1. Number of recruiters by service (same as above)
2. Number of high school graduates by sex (same as above)
3. Number of high school seniors by sex (same as above)
4. Enlistment options available under the Multiple Option
Recruiting Experiment (1979) offered by the Army, Navy,
and Marines

IV. Variables Merged into AFEES and NLS Records at the State level

1. Deviation from trend in total employment (monthly data from
January 1978 to December 1979)
2. Deviation from trend in manufacturing employment (same as
above)

merged in when they become available. Moreover, both the NLS and AFEES samples can be thought of as the first waves of longitudinal surveys. The NLS was explicitly designed as a five-year longitudinal survey, and the 1980, 1981, and 1982 follow-up surveys have already been completed.

Although the AFEES survey was not planned for a follow-up, it can be updated to be longitudinal by linking its records to the personnel records kept by the Defense Manpower Data Center (DMDC). The AFEES survey records have already been linked to DMDC records reflecting outcomes during the first year of service. Data on succeeding years can be included as they become available, and will show how the individuals progressed through their military careers. Similarly, later waves of the NLS can be merged in to provide information on subsequent behavior of the 1979 nonenlistees (e.g., did they enlist at a later date? did they experience job changes?).

One of the file's notable strengths is that it can be enhanced and updated with new information. Without that characteristic, the AFEES-NLS pooled file would be limited to a few studies and then outlive its usefulness. This flexibility gives the file a longer life. It provides the eventual capability to study a panel of individuals through time, and so creates an opportunity to relate a wide range of information available initially, as well as interim information, to subsequent outcomes. Such information has proved useful in studies of employment patterns, earnings growth, labor supply, and migration.

III. USING THE DATA

ADVANTAGES OF MICRODATA

The microdata in the AFEES-NLS database have several advantages for studying enlistment behavior and its policy implications:

- o The richness of such data for comparing populations of enlistees and nonenlistees and for estimating multivariate models of enlistment behavior for different subgroups of young people
- o The opportunity to estimate behavioral equations that are not confused by aggregation bias
- o The ability to control for selectivity bias in analyzing enlistment behavior in conjunction with post-enlistment outcomes, including attrition, promotion, and reenlistment behavior among military personnel.

Richness

As is typical of most microdata, the AFEES-NLS database contains variables not routinely collected nor employed in macro studies of enlistment. As we indicated in Section II, the AFEES-NLS database covers individual demographic characteristics, socioeconomic background, current employment status and wage, work experience, earnings, student status, local recruiters, and local economic conditions. Because the database contains information on nonenlistees, comparisons between enlistees and nonenlistees can easily be made. Also, since the data are suitable for estimating multivariate models, we can analyze whether enlistment behavior differs among different groups of the population or in different areas of the country. For instance, what are the prospects for enlisting more high school degree graduates from the group that is working? Or from the group that has chosen to enter a post-secondary institution?

Avoiding Aggregation Bias

Aggregation bias can occur when equations derived from models of individual choice are estimated with aggregate data. Use of microdata avoids this. The use of aggregate data implies that an explanatory variable's value, and therefore its effect on enlistment, is the same for all individuals. Aggregation bias can arise when individual values differ from the aggregate; the extent of the bias depends on how much the individual and aggregate values differ and on how the variable correlates with the other explanatory variables. The degree of concern about bias depends on whether the models estimated on macrodata are meant to be interpreted as representing micro behavior. For example, macro models typically estimate how wage rates affect enlistments using data on the average wage in manufacturing or on the average wage of young men. But that may not be a good proxy for the wage facing a given young man. Further, its usefulness as a proxy may change over time, because the wages of young men do not necessarily move in lockstep with the average wage in manufacturing. In fact, we expect the relationship between manufacturing wage and youth wage to change cyclically as economic conditions change and secularly as the sizes of youth cohorts change. To our knowledge, empirical models based on aggregate enlistment data have not been rigorously built up from theories of individual behavior, hence raising the question of whether to interpret the coefficients in the macro models with respect to individual behavior. Thus, even though they might lead to accurate forecasts, the aggregate-data models may not be useful in learning, say, how an individual's wage affects his propensity to enlist. But the AFEES-NLS microdata will be useful to do so, and their usefulness will increase as the panels age and individual behavior is observed through time.

Avoiding Selectivity Bias

The AFEES-NLS database, when merged with subsequent service records of those in the AFEES sample, will allow us to study post-enlistment outcomes and to control for selection effects in doing so. This kind of control may be important in predicting the success or failure of entering personnel for attrition, promotion, and reenlistment, and in

analyzing whether their military experience and training exert a positive effect on post-service employment and earnings. Without correcting for selection effects, the coefficient estimates of regressions explaining post-enlistment outcomes will in general be biased. To illustrate, a variable that increases the probability of enlistment and reduces the chance of attrition will have a biased coefficient for attrition because of the enlistment effect. Those who had a high probability of enlisting are, logically, less prone to attrition than those who had a low probability of enlisting; so the coefficient will be estimated to have a value smaller (more negative) than its true value. Thus, the true effect of the variable on attrition can be established only after controlling for its effect on enlistment.

MODELS OF ENLISTMENT CHOICE

Because the project's research on the AFEES-NLS sample will initially concentrate on models of enlistment choice, we can appropriately use that aspect of the research to illustrate what kinds of analyses the data can support. We consider six different models of enlistment choice:

1. Dichotomous choice: enlist, not enlist
2. Service choice: Army, Navy, Marines, Air Force, not enlist
3. Dichotomous choice among stratified sample
 - a. Students: enlist, not enlist
 - b. Nonstudents: enlist, not enlist
4. Polytomous choice among stratified sample
 - a. Students: four services, not enlist
 - b. Nonstudents: four services, not enlist
5. Polytomous choice given enlistment
 - a. Students: four services
 - b. Nonstudents: four services
6. Enlistment and post-enlistment outcomes
 - a. Enlistment
 - b. Attrition from training
 - c. Post-training attrition
 - d. Promotion
 - e. Reenlistment
 - f. Post-service outcomes

The first four models are closely related. The enlistment decision is first defined simply as enlist or not enlist, then elaborated by replacing "enlist" with the choice of entering one of the four services. The models allow for the possibility that enlistment behavior may differ between youths who are students and those who are not. Our preliminary empirical work shows this to be an important distinction, and it supports the opinion of many recruiters that students and nonstudents represent different segments of the recruiting market. Of course, students themselves may differ depending on whether they are in high school or post-secondary school, and full-time or part-time. The AFEES-NLS sample allows us to make these distinctions.

Model 5, choice of service given enlistment, requires only AFEES survey data for its estimation. Models 1 through 4 implicitly or explicitly treat the choice of service as a decision made simultaneously with the choice to enlist. Model 5 assumes that the choice of service occurs sequentially after the choice to enlist. Empirical analysis can be done on models 2 or 4 to determine if that assumption is justified. If the assumption is valid, then a special advantage of model 5 can be exploited: any variable available on the AFEES survey can be used, not only those which the AFEES sample has in common with the NLS sample. Thus the list of potential explanatory variables is longer than in the AFEES-NLS choice-based sample.

Model 6, perhaps the most complex of the models, involves specifying a family of structural equations that describe outcomes of interest: enlistment, training attrition, post-training attrition, promotion, reenlistment, and others. Model 6 can be thought of as a framework for investigating the relationship between information available at the time of enlistment and post-enlistment outcomes. Most work on post-enlistment outcomes so far has concentrated on two variables, high school graduation status and AFQT (or ASVAB) score, for which values are known before enlistment. The AFEES-NLS database can be used to extend the scope of understanding by determining whether other "pre-enlistment" variables also bear importantly on post-enlistment outcomes. Further, model 6 generalizes the approach used in much of the

research on these outcomes by permitting the error terms across the equations to be correlated. The possibility of nonzero error covariance enables the analyst to control and test for selectivity biases in the post-enlistment outcome equations. (See "Avoiding Selectivity Bias" above.) By quantifying the covariance structure that gives rise to selectivity bias, model 6 can be valuable in analyses of alternative accession policies that include eligibility criteria and accession goals by quality. For example, a large increase in accession goals could result in an influx of personnel who, although they might meet all eligibility criteria, have a lower intrinsic attachment to military service. They may be more likely to attrite and perhaps may have lower proficiency in their military skills. The multi-equation model aids in controlling for the unobserved factors associated with such behavior. Moreover, the model can in principle predict how post-enlistment outcomes such as attrition or promotion would vary as accession policies change, other things equal. Existing models of attrition, promotion, and retention cannot do so.

IV. POLICY APPLICATIONS AND RESEARCH TOPICS

POLICY APPLICATIONS

Analysis of the AFEES-NLS database will have relevance for several areas of military manpower policy. For recruiting policy, it can identify the "best" prospects for recruiters to help them meet their enlistment goals, where "best" can be portrayed in terms of the individual's probability to enlist and, as future data become available, his likelihood of succeeding during military service. Research will show the effects of demographic characteristics, socioeconomic status, labor force experience, and local employment opportunities on these outcomes. Furthermore, equations can be estimated for different groups to determine whether behavior differs by student status, age, educational attainment, and so forth. (Such analyses could not be done with data from service personnel files alone, even if the data were made available on a micro basis, because those files do not contain records on nonenlistees.) In this way, studies based on the AFEES-NLS will provide recruiters with specific guidance on the profiles of likely enlistees. This may help steer the recruiter away from prospects who have the right credentials (in terms of educational attainment, and perhaps AFQT score) but who may nevertheless have low predicted chances of entry or of successful first-term performance.

Analyses of the microdata file will also help policymakers prepare for the possible effects that reversals in economic or demographic trends can have on enlistment decisions. For example, enlistment equations will show how an individual's enlistment probability differs across areas facing different industrial mixes and different degrees of cyclical unemployment. Potentially, this information will be useful to predict how enlistments would change if conditions had been better or worse. As another example, suppose the wages and employment opportunities of young men rise over the 1980s as their supply relative to prime age (35-54) men declines. This event would reverse the trend of the 1970s, when the relative supply of youth increased throughout the decade and their earnings and employment fell. AFEES-NLS enlistment models can be employed to predict the consequences of the reversal.

RESEARCH TOPICS THE DATA SUPPORT

Given the panel-like nature of the AFEES-NLS sample, a variety of enlistment studies can be undertaken, some now and others in the near future as data for more years become available.

1. The data can be used to study enlistment decisionmaking during the period when the data were collected, spring 1979. Various enlistment models have been described above and could be estimated for women as well as men, although it must be recognized that women were (and are) subject to different accession policies than men (e.g., high school completion was then required, separate accession goals exist). Current work with the AFEES-NLS sample has considered only young men.

2. The AFEES-NLS data offer a good foundation for comparing intentions to enlist with actual enlistment behavior. These comparisons can proceed in two ways. First, in the 1979 wave, the NLS asks about the individual's intention to enlist. That information can be correlated with the individual's actual enlistment behavior as recorded in the 1980 or subsequent NLS waves. This method has a serious limitation: a small number of persons in the sample will have enlisted by 1980. About 3100 of the roughly 5000 non-prior-service young men in the NLS were available for enlistment in 1979 (i.e., old enough and without a health condition that would prevent their working full-time). Based on recent experience, perhaps 25 to 30 percent of these 3100 young men will ever serve on active duty, and the vast majority of those will have entered by their early twenties. The AFEES portion of the AFEES-NLS database, in contrast, contains about 5300 enlistees. Alternatively, we can compare the tables or regressions characterizing enlistment intentions in the NLS with tables or regressions describing actual enlistment behavior, as estimated from the AFEES-NLS sample. But this method does not allow us to compare the individual's enlistment intentions with his own subsequent enlistment behavior.

3. Enlistment behavior observed over time in the NLS can be compared with cross-sectional enlistment behavior as estimated from the AFEES-NLS data. AFEES-NLS regression models can, as mentioned, be estimated for different groups of the population. For instance, the group of 20 year olds not in school consists of persons who for several

years after high school chose not to enter the military. The AFEES-NLS allows us to control for labor force experience within the past year; the NLS alone, as more years of data accumulate, will permit control for labor force experience over a longer period. We can thus learn whether additional years of data on labor force history (or other personal history, e.g., marriage, education, health) contribute to the explanation of enlistment behavior.

4. The AFEES data, when linked with DMDC personnel records, can be employed to study attrition, promotion, and retention behavior. Because of their abundant detail, the AFEES data permit greater control for individual background characteristics in multivariate analyses; previous studies, limited to variables available on service personnel records, have only been able to control for a few such characteristics.

5. The NLS data can be used in parallel with the AFEES survey data to study school and labor market experience of young men and women. For example, the job turnover experience of young men and women in the private sector can be compared with that of young men and women in the military. The AFEES-NLS database facilitates these comparisons by creating variables that are virtually equivalent for the two surveys, and by supplementing each observation with data extracted from other files.

SOME ENLISTMENT HYPOTHESES

The microdata file can be used to test hypotheses about enlistment and post-enlistment behavior. A fundamental hypothesis of enlistment behavior states that individuals are more likely to enlist the higher the military wage is relative to the civilian wage. In aggregate models, the hypothesis has been tested by seeing whether the ratio between military and manufacturing wages has a positive effect on enlistments. In micro models based on the AFEES-NLS data, we can specify the hypothesis in terms of the actual wage the individual receives on his current job, or if he is not employed, the wage he received on his most recent job. Because wage variation across individuals is ample, the data present a good opportunity to estimate how higher wage rates affect enlistment. Moreover, the wage variable can be interacted with the person's student or nonstudent status. We

would expect wage to measure the price of time more meaningfully for a nonstudent than a student: many students hold part-time jobs to help meet expenses, but the jobs may be completely unrelated to their skills or interests.

Another basic hypothesis holds that enlistments are related to changes in unemployment. From aggregate models we know that higher unemployment rates lead to increases in enlistments, but we do not know the pathway of the response. An increase in the unemployment rate may signal an increase in the extent of unemployment among young men, and consequently they may be more likely to enlist. But higher unemployment might also mean a decline in current or expected future private sector earnings, so that even employed young men could find enlistment more attractive. In addition, a connection could exist between the individual's employment status and local economic conditions: unemployed people who live in a boom area may be less likely to enter the military than unemployed people in a depressed area. Each of these hypotheses can be tested with the AFEES-NLS data, and thereby supplement our understanding of why enlistments are related to unemployment in the aggregate.

Some brief examples of other hypotheses that can be tested with the AFEES-NLS database are:

- o An individual's labor force experience affects his propensity to enlist. Perhaps older enlistees tend to come from the pool of persons who have not fared well in the labor market.
- o The influence of educational expectations on the enlistment decision may vary considerably among men in high school, in post-secondary school, or in the labor force. This could signal a differential enlistment responsiveness to advertising or enlistment incentives related to a college education.
- o The socioeconomic status of the parents influences enlistment. Like the wage and education-desired variables, the socioeconomic variables may inhibit the social representativeness of the AVF.

Each hypothesis can be analyzed by data that have been stratified by variables such as age, sex, education status (full-time student; non-full-time student), and AFQT category. The stratifications can identify whether the different groups behave differently and so might be viewed as separate segments of the recruiting market. Finally, most of these hypotheses can be repeated in terms of post-enlistment behavior as the panel ages.

V. CONCLUSIONS

The AFEES-NLS database will be useful for investigating different models of enlistment and post-enlistment outcomes. Moreover, these microdata expand the range of testable hypotheses well beyond that already covered by aggregate data. As explained, the richness of the AFEES-NLS database derives not only from the range of variables available for analysis, but also from its longitudinal nature. We expect the data to help quantify differences in the probability of enlistment across different segments of the enlistment pool. At the same time, the data will show whether the segments respond differently to changes in employment conditions, military/civilian pay, and other explanatory variables. Further, by studying enlistment and post-enlistment outcomes as a system of equations the AFEES-NLS data can help predict first-term performance as a function of information available at the time of entry--education and ASVAB scores, as well as sociodemographic and labor force experience variables. On net, the AFEES-NLS database can offer new information about who is more or less likely to enter the military, which service he will choose, whether he is likely to make it through the first term, and how these outcomes would vary according to changes in the variables influencing the decision to enlist.

Should such information prove useful to OSD and the services, the methodology employed in creating the AFEES-NLS database can be replicated. In particular, OSD surveys of enlistees can again be coordinated with other surveys involving youth; this would pave the way for constructing additional choice-based samples. Ongoing surveys involving youth include the monthly Current Population Survey, which uses a rotating sample (with the sequence, on four months, off eight, on four); the 1979 High School and Beyond survey, which is panel in nature; and others. In addition, by consciously planning defense survey efforts to dovetail with future survey efforts by nondefense agencies, choice-based sampling will make both survey efforts more cost-effective than previously possible.

Appendix
CREATION OF THE AFEES-NLS DATABASE

The creation of a choice-based sample through pooling separate data files requires data collection from the same time period and comparable information across files. Our microdata file satisfies those two criteria. The sample of enlistees comes from the first wave of the 1979 DoD Survey of Personnel Entering Military Service (referred to as the AFEES survey) and the sample of nonenlistees comes from the 1979 wave of the new National Longitudinal Survey of Labor Force Behavior, Youth Survey (referred to as the NLS survey). Each survey was administered in the spring of 1979 and concentrated on demographics, education, and employment histories. In addition, both contained geographic indicators which would enable us to add data on local employment conditions, youth population, recruiter population, and availability of enlistment options.

The 1979 NLS youth survey was administered to a stratified random sample of young men and women aged 14 to 22 in 1979. The majority of interviews, about 63 percent, were conducted in March, April, and May. We recognize that not all youth in the NLS file were eligible for enlistment in 1979. The survey includes those who were too young, those with physical handicaps, and a supplemental sample of 800 young men and 400 young women who were already serving in the armed forces. Also, some youth will be ineligible because of low scores on the Armed Forces Qualifying Test; however, eligibility based on AFQT results cannot be determined until AFQT scores become available for the NLS observations later this year. (The 1981 wave of the NLS will include 1980 ASVAB test scores.) Individuals determined to be ineligible for enlistment in 1979 due to age or health conditions were deleted when the analysis file was constructed.

The 1979 AFEES survey was administered at all 67 AFEES (Armed Forces Entrance Examination Stations) to those non-prior-service individuals who had signed enlistment contracts in April-May and September-October of that year. Two questionnaires were used in each

wave. We selected only the second questionnaire from the first wave to pool with the 1979 NLS youth survey. The first wave was chosen because of the proximity of interview dates, and the second questionnaire was selected because of the larger number of variables similar to those on the NLS questionnaire. The geographic identifier on the AFEES for each respondent's residence is the zip code. By means of a cross reference program, the zip codes were mapped into FIPS (Federal Information Processing Standards) county codes, the same kind of geographic identifier appearing on NLS records.

Table A.1 shows the overlap in interview dates between the 1979 AFEES and NLS samples.

Table A.1
CORRESPONDENCE OF INTERVIEW DATES BETWEEN 1979 AFEES AND 1979 NLS
(Percent)

| Interview Month | AFEES | | ^a NLS | |
|--------------------|-------|---------|---------------------|---------|
| | Males | Females | Males | Females |
| Jan | 0.1 | 0.2 | 0.0 | 0.1 |
| Feb | 0.0 | 0.1 | 24.0 | 26.7 |
| Mar | 0.4 | 0.4 | 44.0 | 40.4 |
| Apr | 34.4 | 33.2 | 19.6 | 20.4 |
| May | 63.3 | 64.6 | 8.9 | 9.9 |
| June | 1.5 | 1.3 | 1.4 | 0.7 |
| July | 0.0 | 0.0 | 1.8 | 1.5 |
| Aug | 0.2 | 0.2 | 0.3 | 0.3 |
| Sept- | 0.1 | 0.1 | 0.0 | 0.0 |
| Dec | -- | -- | -- | -- |
| Total | 100 | 100 | 100 | 100 |
| N | 5252 | 1405 | 3125 | 3494 |

^a
Sample is non-prior-service, age 17 and over, no health conditions that impede ability to work.

In pooling the two files, all the AFEES (Wave 1, Form 2) questions were retained while only a subset of the NLS's 6000 plus variables were kept. We extracted from the NLS all the questions we thought would be necessary to create the comparable variables listed in Table A.2, an abbreviated version of Table 1.

If additional information is ever needed, we can return to the NLS, extract the desired variables, and merge them with the AFEES-NLS. Exhibit 1 lists the questions extracted from the NLS youth survey.

From this basic pooled file, we created a subsequent file with identical records for AFEES and NLS observations.

We then merged in geographic-based data to provide some measures of the environment in which the decisionmaking process occurred. Certain data were unavailable for 1979 and thus we added the most recent data we could find. For population data this may not pose a serious problem since relative sizes do not change much. However, for economic data, it should be noted that, say, 1977 employment data are not an appropriate indicator of economic conditions in 1979. Past economic data can be used, though, to look at past trends within the geographic unit.

Table A.2

COMPARABLE VARIABLES BETWEEN 1979 AFEES AND 1979 NLS SURVEYS^a

1. Enlistment status
2. Age
3. Race
4. Education
5. ASVAB scores (available fall 1982)
6. AFQT percentile (available fall 1982)
7. Marital status
8. Socioeconomic background
9. Labor force experience last year (1978)
10. Earnings experience
11. Current employment experience
12. Location of residence

^aFor a more complete description, see Table 1.

Data were merged in at three geographic levels: county, AFEES, and state. (In this context, AFEES refers to the Armed Forces Entrance Examination Station, not the AFEES sample.) County-AFEES mapping was used to merge AFEES-level information into the NLS, whose observations have county level designators. The county-level data consist of the following:

1. Number of recruiters by service (spring 1979)
2. Number of high school graduates by sex (1977 data)
3. Number of high school seniors by sex (1978 data)
4. Population counts by sex and race (white/black-minority) for age groups 10-14, 15-19, 20-24, 25-29 (1978 estimates)
5. Employment levels by one-digit SIC industry, annually, for the years 1969 to 1977

AFEES-level data include the following:

1. Number of recruiters by service (same as above)
2. Number of high school graduates by sex (same as above)
3. Number of high school seniors by sex (same as above)
4. Enlistment options available under the Multiple Option Recruiting Experiment (1979) offered by the Army, Navy, and Marines

State-level data consisted of the following:

1. Deviation from trend for total employment for each month (January 1978 to December 1979)
2. Deviation from trend for manufacturing employment for each month (January 1978 to December 1979)

Exhibit 2 contains source references and brief descriptions for the above data. Exhibit 3 lists the enlistment options available under the Multiple Option Recruiting experiment.

Lastly, for the AFEES survey data we were able to map in information from the individual's Military Enlistment Processing Command (MEPCOM) record, which is maintained on file by the Defense Manpower Data Center. The MEPCOM record contains data not available on our AFEES questionnaire but useful in analysis, such as AFQT, MOS (Military Occupation Specialty), enlistment options taken, etc.

The above process describes the creation of the basic AFEES-NLS file. At this stage no deletions were made and no variables comparable across files were created. The analysis file was created from this file.

In addition to the choice-based weighting required in multivariate analyses of enlistment, both the NLS and AFEES observations need to be weighted separately. The NLS is a stratified random sample and therefore must be weighted to reflect the population at large. Sample weights for each observation exist on the file. The AFEES, while not a stratified random sample, may require weights to correct for nonresponse rates. The questionnaire was to be administered to all enlistees who came in during the survey period. However, response rates were less than 100 percent; thus the AFEES observations may need to be weighted to reflect the population who enlisted in spring 1979. Work is under way to determine the need for weighting and to develop weights if necessary.

COMPARABILITY ISSUES IN THE CREATION OF AN ANALYSIS FILE

The crux of creating a choice-based sample from pooling different surveys is comparability of data. Variables must measure the same thing across all observations regardless of the choice made. If not, errors or biases may be introduced which lead to coefficient estimates and significance tests that are misleading.

This subsection describes difficulties which arose during the process of defining variables that were comparable across files. In the simplest cases, it was enough to redefine variable codes to make them identical across files; responses were just rearranged in categories defined equivalently for AFEES and NLS observations. However, for many variables this procedure was not sufficient--either additional information was needed or a whole new variable had to be constructed from other variables available in the file.

Although the NLS and AFEES ask similar types of questions, they are by no means exact. A simple question like "years of schooling completed" appears on both surveys, but the NLS refers to years completed up to the survey and the AFEES refers to anticipated years completed by the time the respondent enters active duty. What results,

at least for students, is a mean difference of one year of schooling between the two files--the NLS respondents have generally completed one year less of school than the AFEES respondents. To create an NLS variable comparable to the AFEES variable we used the year of school being attended at the time of survey and assumed the respondent completed that year.

The AFEES survey tended to ask simpler, more direct questions than did the NLS. For instance, the NLS went into considerably greater depth about many aspects of the youth's work experience, both past and present. Therefore, to create AFEES-comparable employment variables for NLS observations, we often have to piece together numerous bits of information to form one variable. A good illustration of this concerns the AFEES question, "What were you doing most of the time each month from January 1978 to February 1979?" (Q55A-Q55N in Wave 1, Form 2). The respondent could chose among four possible responses for each month: worked full-time, worked part-time, unemployed and looking for work, or none of these. (Moreover, the criteria for full- and part-time work were left to the respondent to decide.) There are no direct counterparts to these responses in the NLS survey; they had to be created from information on previous jobs, lengths of employment, periods of unemployment, and so on. It was also necessary to specify the criteria for full- and part-time work for the individual. As a result, the NLS observations now have a nearly, but not perfectly, comparable version of the AFEES question.

Questions that appear similar may still differ because of different skip patterns between the two surveys, and for this reason we had to be careful to define variables for the same subset of people in each file. "Reason left school" and "part-time, full-time student" are typical examples. In the AFEES only those who do not have at least a high school diploma are asked why they left school (i.e., high school), and student status is asked of everyone. In contrast, the NLS asks anyone who is not currently attending school why he left school. This group includes both graduates and dropouts since a valid reason for leaving school is graduation. Also, student status is asked only of those who are attending or have attended college.

A difference in emphasis can cause seemingly similar variables to be quite different. The definition of family income is a classic example. In the AFEES survey, family income refers to the respondent's parents' income, regardless of whether he resides with them. The NLS, on the other hand, equates family income with household income; this is the same as parents' income only when the respondent still lives at home. Thus, family income can be applied equitably only to the subset of respondents in both surveys who live with their parents. Because of this restriction, father's and mother's education may be better proxies for family (i.e., parents') income than the family income variable itself.

Another difficulty in achieving comparability arises when the surveys have different objectives in asking a particular question, so the manner in which the question is asked can affect the similarity of the surveys. Current job information, for example, refers to something slightly different in each file. The AFEES asks about the job the respondent has at the time of the survey, or, if the respondent is unemployed, about his last (most recent) job. The NLS asks what the respondent was doing most of last week; if he was working, that job becomes his current job (or his last job if he is not employed at survey time). If he was not working last week there is no current job information, and one must look to the "second job" listed to get information on the last job held. The NLS asks questions about the current job in this way so they will correspond to questions in the Current Population Survey on current employment. In any event, this subtle difference between the AFEES and NLS approaches caused us trouble in trying to arrive at comparable information about the respondent's current job or last job held.

How each survey was administered can also affect the comparability process. The NLS was conducted with an interviewer who could cross check responses to ensure consistency in the data--if a youth said he was not married, no marriage questions were asked. The AFEES survey, on the other hand, was self-administered, and there was no supervision to see if the proper sections were skipped or answered. Further, when the AFEES survey questionnaires were processed, inconsistent responses were

not altered. If the respondent said he never had a job but filled in questions on current/last job, the responses were allowed to stand. In developing our AFEES-NLS analysis files we had to decide which responses were valid.

The complicated nature of the NLS survey questionnaire made the process of reaching AFEES-NLS comparability tedious. Each variable was examined for the problems mentioned above, and the comparisons and checks continued throughout the preliminary data analysis. Odd results generally proved traceable to inadvertent comparability miscodings. Fortunately, the ability to return to the basic file enabled us to correct the errors.

Information on how specific variables were made comparable across surveys may be obtained from the authors. The level of detail necessary to describe the comparability process for each variable is beyond the scope of this Note and has, therefore, been mentioned in general terms.

Exhibit 4 gives the variable list of the analysis file where comparable variables were created and those ineligible for enlistment were dropped.

Exhibit 1

VARIABLES EXTRACTED FROM 1979 NLS SURVEY

Definition

1. Age of respondent
2. Who lived with at age 14
3. Education of mother
4. Education of father
5. Number of siblings
6. Number of older siblings
7. Education of eldest sibling
8. Ethnic origin(s)
9. Ethnic origin most identify with
10. Marital status
11. Month and year of first marriage
12. Month and year first marriage ended
13. Month last marriage began
14. Year last marriage began
15. Month and year of most recent divorce
16. Month and year present separation began
17. Number of children now
18. Number of children expected
19. Currently enrolled
20. Month and year last attended school
21. Reason not currently in school
22. Highest grade attended
23. Highest grade completed
24. Respondent has high school diploma/GED
25. Has diploma or has GED
26. Full-time or part-time student (college)
27. Highest grade would like to complete
28. Highest grade expect to complete
29. Expect to enlist in future
30. Number of hours worked during survey week
31. Total number of hours worked during survey week
32. Number of weeks since current layoff began
33. Number of weeks looking for work (new job to begin)
34. Number of weeks looking for work (unemployed)
35. Have any other job since January 1978 (employed)
36. Have any job since January 1978 (unemployed or out of the labor force)
37. Had any regular job since January 1978
38. Currently working at job 1,2,3,4 or 5 (employer flap)
39. Reason left job 1
40. Reason left job 2
41. Hours per week worked--jobs 1,2,3,4 and 5
42. Government sponsored job (jobs 2,3,4 and 5)
43. Respondent's age at jobs 2,3,4 and 5
44. Worked < or \geq 20 hours per week at jobs 2,3,4 and 5

45. Did job begin before Jan 1978 (jobs 1,2,3,4 and 5)
46. Any periods not working at jobs 1,2,3,4 or 5 since Jan 1978
47. Rate of pay job 1
48. Time unit of pay for job 1
49. Any vocational or technical training since Jan 1978
50. Completed vocational training program: 1st,2nd,3rd program
51. Any vocational/technical training before Jan 1978
52. Completed vocational training taken before Jan 1978
53. Week # of end of 1st-6th period of no work since Jan 1978
54. Week # of start of 1st-6th period no work since Jan 1978
55. Could health prevent working for pay now
56. Does health limit kind of work can do
57. Does health limit amount of work can do
58. Name of main health condition that limits work
59. Personal income 1978
60. Race of respondent
61. Sample I.D. code
62. Screener Case I.D. number
63. Total net family income 1978
64. Respondent's sex
65. Health conditions that limit work (5-digit codes)
66. Family income 1978 (truncated) of person in item 18
67. Total net family income 1978 (truncated)--household record
68. Total net family income 1978 of resp. and family (truncated)
69. Number of weeks worked since 1978 (ages 16+)
70. Number of weeks unemployed since 1978 (ages 16+)
71. Number of weeks out of labor force since 1978 (ages 16+)
72. Number of weeks since 1978 (ages 16+)
73. Number of weeks worked in 1978 (ages 16+)
74. Number of weeks unemployed in 1978 (ages 16+)
75. Number of weeks out of labor force in 1978 (ages 16+)
76. Month, day, year began work at jobs 1,2,3,4 and 5
77. Month, day, year ended work at jobs 1,2,3,4 and 5
78. Month of interview
79. Sampling weight version III
80. Ever served in armed forces

Exhibit 2

REFERENCES FOR GEOGRAPHIC DATA

1. Multiple Option Recruiting Experiment enlistment options by AFEES

Description: In 1979, the Army, Navy and Marine Corps instituted experiments of varying available enlistment options across the country. The basic options offered were a 2-year term of enlistment and VEAP (Veterans' Educational Assistance Program) "kickers" (lump sum additions to an individual's VEAP funds based on \$1000 to \$2000 per year of commitment). These options were generally offered to "high-quality" (Cat I-III A) persons enlisting in designated occupations. For a more detailed description of the options available see G.W. Haggstrom, T.J. Blaschke, W. Chow, and W.J. Lisowski, The Multiple Option Recruiting Experiment, R-2671-MRAL, The Rand Corporation, November 1981.

Source: Army, Navy, Marine Corps

2. Recruiter population by FIPS county code and AFEES station code

Description: Figures are based on the number of "effective" recruiters as agreed upon by the four services to standardize for full- and part-time recruiters employed at different rates among the services. A Joint Marketing and Advertising Research algorithm was used to apportion recruiters among counties within the AFEES. The data pertain to April 1979.

Source: Recruit Market Network, DMDC

3. High school graduates and seniors by FIPS county code and AFEES code

Description: In 1976, the Curriculum Information Center in Denver, Colorado, conducted a complete telephone survey of all high schools in the United States. The figures for 1977 are estimates based on a 5 percent sample survey used to update the 1976 data. The estimated number of high school graduates is for 1977; the estimated number of high school seniors is for the 1977-1978 school year.

Source: Recruit Market Network, DMDC

4. Youth population by age, sex and race by FIPS county code

Description: The Bureau of the Census prepared experimental population estimates by age, sex, and race for the years 1970-1978 for all U.S. counties for use by the National Cancer Institute in 1975. The estimates were constructed by carrying forward the 1970 Census counts by age cohort using registered births and estimated deaths by age and basing the age-sex detail of migrants on the information provided for the period 1965-1970 by the 1970 Census. County race detail is not very accurate; it is an extension of the 1965-1970 migration trends which are used to adjust the state averages to create county-specific figures. For more detailed information see: Current Population Reports, Special Studies, P-23, No. 103, "Methodology for Experimental Estimates of the Population of Counties by Age and Sex: 1975"

Source: U.S. Department of Commerce, Bureau of the Census

5. Employment levels by 1-digit industry by FIPS county code

Description: The Bureau of Economic Analysis tallies employment levels by 1-digit SIC codes for all counties in the U.S. The data are taken from ES-202 forms which employers must fill out specifying their number of employees and their employees' earnings. The Bureau adjusts the figures for small firms (one to four employees) for their lower response rates. Information for proprietors is acquired from the Internal Revenue Service. The data run from 1969 to 1977.

Source: Bureau of Economic Analysis, State and County Employment and Earnings Division

6. Deviation from trend in employment by state code

Description: Time trend regressions using monthly state employment figures for the period 1976-1981 were run for each state separately. The residuals from those regressions are used as measures of employment variability (deviations from trend) corrected for seasonality (e.g., a positive residual indicates employment levels above the norm for that month). The employment figures came from State and Area Employment, Hours and Earnings produced by the Bureau of Labor Statistics. The data are taken from payroll records submitted voluntarily to state agencies by a representative sample of industrial, commercial, and government establishments employing, collectively, about 32 million workers. Residuals for each

month from January 1978 to December 1979 were added to the file. Data from previous or future periods may be added if necessary.

Source: Bureau of Labor Statistics

Exhibit 3

MULTIPLE OPTION RECRUITING EXPERIMENT ENLISTMENT OPTIONS
AVAILABLE IN SPRING 1979

| <u>Army</u> | <u>Navy</u> | <u>Marines</u> | <u>Air Force</u> |
|---|---|-----------------------|-----------------------|
| <u>VEAP kicker</u> <u>with enlistment</u> | <u>VEAP kicker</u> <u>with enlistment</u> | <u>VEAP kicker</u> | <u>VEAP kicker</u> |
| 2-4 year | 4 year | | |
| 2-4 year, 2 year option | 2 year, "A" school immediately | | |
| 2-4 year, European assign. required | 2 year, "A" school at reenlistment | | |
| 2-4 year, 2 year option, European assign. required | | | |
| 3-4 year, 2 year option, European assign. required | | | |
| 2-4 year, 2 year option, Indiv Ready Reserve | | | |
| <u>No VEAP kicker</u> | <u>No VEAP kicker</u> | <u>No VEAP kicker</u> | <u>No VEAP kicker</u> |
| | 2 year enlistment | 2 year enlistment | |
| | 2 year enlistment, "A" school at reenlistment | | |

NOTE: Option combinations were generally offered only to high-quality (Cat I-III A) individuals enlisting in specific occupations.

Exhibit 4

1979 AFEES/NLS DATABASE VARIABLE LIST (Analysis File,
November 1982 Version)

I. AFEES-NLS Comparable Data

Sex
Age
Race
Who lived with at age 14
Currently enrolled
Last month attended school
Last year attended school
Full-time or part-time student
Highest grade completed
Reason left high school
Highest grade expect to complete
Marital status
Length of current marriage
Number of children
Number of children expected
Starting month of last job
Starting year of last job
Average hours per week worked at last job
Hourly wage at last job
Ending month of last job
Ending year of last job
Currently employed
Tenure of current job
Length of current unemployment
Reason left job - laid off
Reason left job - fired
Reason left job - moved
Reason left job - family reasons
Reason left job - bad working conditions
Reason left job - bad pay
Reason left job - illness
Reason left job - school
Reason left job - other reason
Employment pattern 1978
Personal income 1978
Number of siblings
Number of older siblings
Education of eldest sibling
Education of father
Education of mother
Family income 1978
AFQT Percentile (renormed for 1979 AFEES)
ASVAB test scores

II. AFEES Survey Data

The entire AFEES questionnaire was retained since additional analysis would be done on AFEES observations alone.

Positions 1-643 of source record

III. NLS Data

Most of the NLS variables originally extracted were used to create comparable variables between the AFEES and NLS surveys. Only those variables which might have subsequent use were carried over to the working file. Any additional NLS data that may be required can easily be merged into the working file from the original NLS file.

- (V579) Type of job - second job
- (V580) Type of job - third job
- (V581) Type of job - fourth job
- (V582) Type of job - fifth job
- (V583) Age level - second job
- (V584) Age level - third job
- (V585) Age level - fourth job
- (V586) Age level - fifth job
- (V666-V667) Actual wage or salary - first job
- (V668) Wage base - first job
- (V1041) Vocational training after 1977
- (V1058) Vocational training before 1978
- (V1217-V1218) Personal income 1978
- (V1572-V1573) Family income 1978
- (V2177) Month of interview
- (V2179-V2180) Sample weight

IV. MEPCOM DEP/Accession Data

A subset of the entire MEPCOM record were merged in; only those data that seemed pertinent to the enlistment study were retained.

- Match flag
- Record identification
- Service
- Prior service
- Home of record FIPS
- Age at entry
- Highest year of education
- Sex
- Race
- Marital status/dependents
- Test form
- AFQT percentile (not renormed)
- Mental category

Entry status
Waiver code
Waiver approval level
Date of entry into DEP/date of accession
Term of enlistment
Entry pay grade
Program enlisted for
Bonus option
Enlistment option
Entry MOS

V. Sample and Geographic Identifiers

Sequence number (100001-107332=AFEES, 200001-212686=NLS)
ZIP code
FIPS code
Recruit Market Network AFEES code
MEPCOM AFEES code

VI. Multiple Option Recruiting Experiment (MORE) Data by AFEES Code

Match flag
Army MORE options
Navy MORE options
Marine Corps MORE options

VII. Recruiter/High School Population Data by FIPS County and by AFEES Code

Match flag
Number of recruiters - Army
Number of recruiters - Air Force
Number of recruiters - Marine Corps
Number of recruiters - Navy
Number of high school graduates - male
Number of high school graduates - female
Number of high school seniors - male
Number of high school seniors - female

VIII. County Youth Population Data

Match flag
White male population, ages 10-14
White male population, ages 15-19
White male population, ages 20-24
White male population, ages 25-29
White female population, ages 10-14
White female population, ages 15-19
White female population, ages 20-24
White female population, ages 25-29
Black/minority male population, ages 10-14
Black/minority male population, ages 15-19
Black/minority male population, ages 20-24
Black/minority male population, ages 25-29
Black/minority female population, ages 10-14
Black/minority female population, ages 15-19

Black/minority female population, ages 20-24

Black/minority female population, ages 25-29

IX. County Employment Data by 1-Digit (SIC) Industry

Match flag

1969-77 employment - full-time and part-time

1969-77 employment - hotels

1969-77 employment - entertainment and recreation

1969-77 employment - miscellaneous business and repair

1969-77 employment - business

1969-77 employment - agriculture, forestry, fish and game

1969-77 employment - mining

1969-77 employment - construction

1969-77 employment - trans, comm, public util

1969-77 employment - wholesale

1969-77 employment - retail

1969-77 employment - finance, insurance and real estate

1969-77 employment - public administration

1969-77 employment - manufacturing/durable

1969-77 employment - manufacturing/nondurable

1969-77 employment - proprietors

1969-77 employment - proprietors/farming

1969-77 employment - personal and private

1969-77 employment - professional and related

X. Deviations from trend by state (employment variability)

Monthly residuals from time trend regression on monthly state
employment rates

Total employment residual by month for 1978 and 1979

Manufacturing employment residual by month for 1978 and 1979

END

FILMED

6-83

DTIC